CIPE
MAY 07 2007
PATENT & TRADEMARK OFFICE
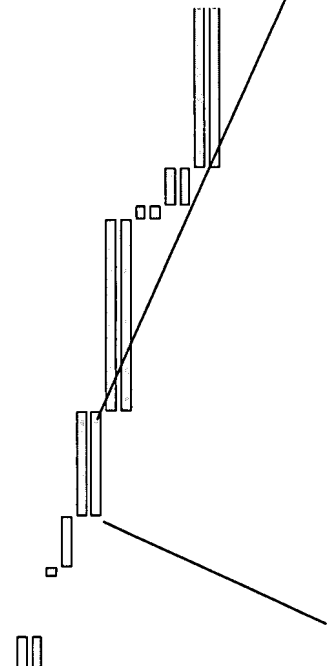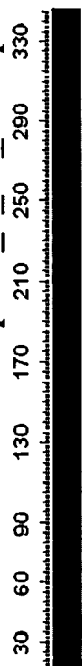
Sample 1: Heteroplasmy (mixture of lengths) is detrimental to sequencing approaches but does not adversely affect MS measurements

HV1-1-outer-variant1: p482_3_26_2003.pek, scan HV1-1-outer-variant2: p482_3_26_2003.pek, scan 1

30 60 90 130 170 210 250 290 330

Scan number: 1
Position: 94 to 145
Length: 52
Number of matched peaks: 1
Fragment seq: CATAAAACCCAATCCACATCAAACCCCCCCCCCATGCTTACAAGCAAGT
Average Predicted fragment mass: 15753.66266 + 0.00000 daltons
Average matched peak mass: 15753.63000 ± 0.00000 daltons
Total peak abundance: 4.01E+06
Average peak abundance: 4.01E+06 ± 0.00E+00
Maximum peak abundance: 4.01E+06
Weighted average match error: 2.07327 ± 0.00000 ppm
Raw average match error: 2.07327 ± 0.00000 ppm
Weighted average peak fit: 0.05900 ± 0.00000
No ambiguous assignments in group

30 60 90 130 170 210 250 290 330

Scan number: 1
Position: 94 to 145
Length: 53
Number of matched peaks: 1
Fragment seq: ACATAAAACCCAATCCACATCAAACCCCCCCCCCCATGCTTACAAGCAAGT
Average Predicted fragment mass: 16042.70903 ± 0.00000 daltons
Average matched peak mass: 16042.68880 ± 0.00000 daltons
Total peak abundance: 9.97E+06
Average peak abundance: 9.97E+06 ± 0.00E+00
Maximum peak abundance: 9.97E+06
Weighted average match error: 1.26121 ± 0.00000 ppm
Raw average match error: 1.26121 ± 0.00000 ppm
Weighted average peak fit: 0.02100 ± 0.00000
No ambiguous assignments in group

(opposite strand also observed)

HV1-1-outer-variant1   acataaaaacccaatccacatcaaacccccccccc atgcttacaagcaagt
HV1-1-outer-variant2   acataaaaacccaatccacatcaaaccccccccccc atgcttacaagcaagt

(opposite strand also observed)   (SEQ ID NO: 44)
(SEQ ID NO: 45)
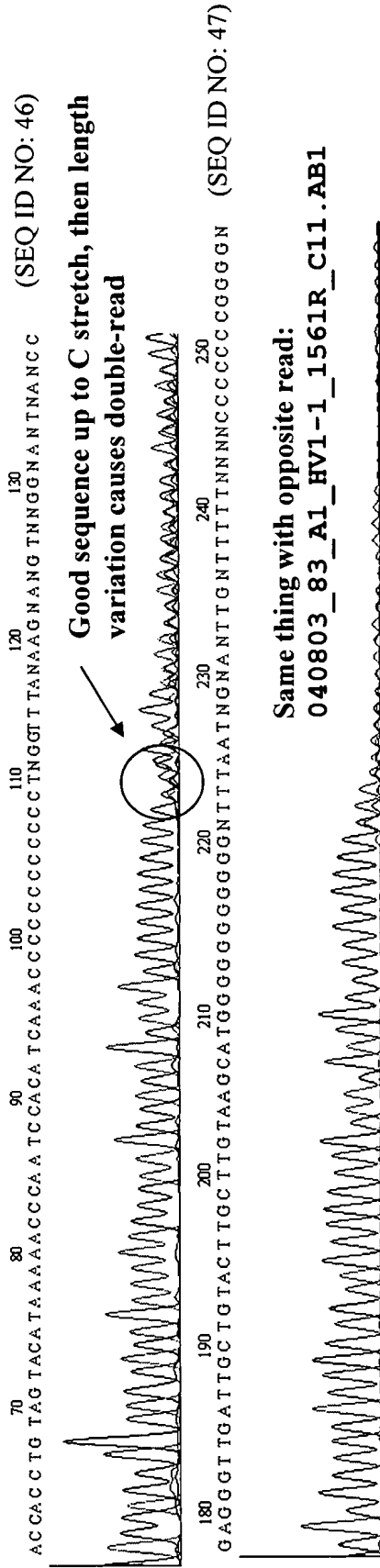
Heteroplasmic region

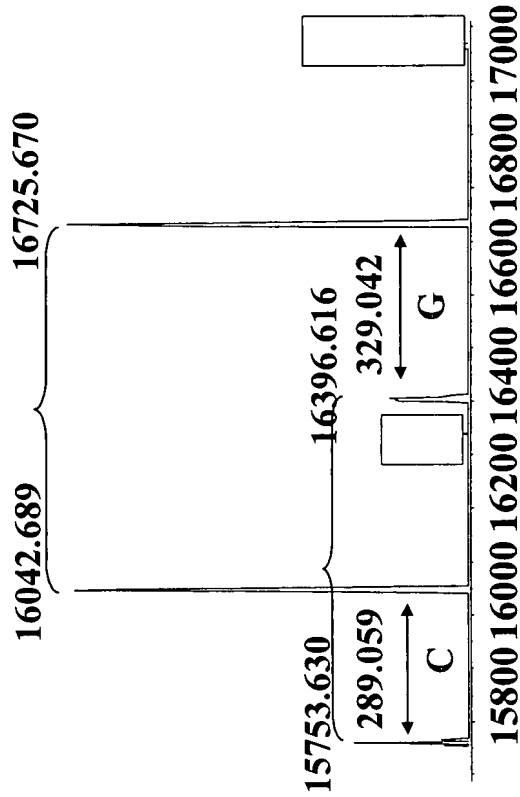### Figure 20A

"Replacement Sheet"

Figure 20B

MS Approach Succeeds Where Conventional Sequencing Fails

Sequences differ in length by one base in the 'C' stretch. Both variants appear in digest data. Sequencing profile also points to a length variation in this region: Example trace: 040803_81_A1_HV1-1_1560F_A11.AB1 (SEQ ID NO: 46)

Good sequence up to C stretch, then length variation causes double-read

ACCACCTG TAG TACATAAAAAACCAA TCCACA TCAAACCCCCCCCCC TNGGTT TANAAGNANG TNNGGNANTNANCC

GAGGGT TGATTGC TGTACTTGC TTGTAAGCATGGGGGGGGGGNTTTAATNGNANTNGNTTTTNNNNCCCCCCGGGGGN (SEQ ID NO: 47)

Same thing with opposite read: 040803_83_A1_HV1-1_1561R_C11.AB1

MS detects multiple species simultaneously while sequencing requires pure sample for maximum information content

Ratio of short to long alleles is 1:3

16725.670

16042.689

16396.616

329.042 G

15753.630

289.059 C

15800 16000 16200 16400 16600 16800 17000